Progress Report to

Joint NOAA and NASA
GEWEX Americas Prediction Project (GAPP)

USE OF CLIMATE FORECASTS IN MULTIPURPOSE RESERVOIR SYSTEM MANAGEMENT

June 5, 2003

Lead-PI: David W. Watkins, Ph.D.
Assistant Professor
Dept. of Civil & Environmental Engineering
Michigan Technological University
Houghton, MI 49931
Tel: (906) 487-1640
Fax: (906) 487-2943
E-mail: dwatkins@mtu.edu

Co-PI: Deborah K. Nykanen, Ph.D.
Assistant Professor
Dept. of Civil Engineering
Minnesota State University, Mankato
Mankato, MN 56001
Tel: (507) 389-6383
E-mail: deborah.nykanen@mnsu.edu

Collaborator: Jobaid Kabir, Ph.D., P.E.
Senior Water Resources Planner
Lower Colorado River Authority
P.O. Box 220
Austin, TX 78767-0220
Tel: (512) 473-4064
E-mail: jobaid.kabir@lcra.org

Duration: September 1, 2002 – August 31, 2004

Progress in Last 9 Months

This study seeks to provide water managers with up-to-date information on the viability of seasonal or annual hydrologic forecasts, and outline a means by which these forecasts can be used operationally. All hydrologic data, results, and conclusions pertain specifically to water supply management in the Lower Colorado River basin in Texas, but the analysis methods are fully transferable to other locations and water management problems. In Year 1 of the project, Steps 1-3 were to be completed as follows:

1) Analysis of potential climate predictors for the case study region in central Texas
2) Derivation of maximal skill forecasts based on identified predictor variables
3) Generation of stream flow ensembles consistent with the skill and uncertainty of the forecasts.

Analysis of Climate Predictors

Several potential climate predictors have been analyzed, including stream flow persistence, soil moisture, Southern Oscillation Index (Jones, 2001), North Atlantic Oscillation (Hurrell, 2002), and Pacific Decadal Oscillation (Mantua, 2001). Only weak correlations (0.2-0.3) between these predictors and stream flow have been found with the lead times needed for interannual water supply planning. These results were at first surprising until two previous studies identifying teleconnections for central Texas were reexamined. First, Piechota and Dracup (1996) found strong correlation between the Southern Oscillation Index (SOI) and the Palmer Drought Severity Index (PDSI). However, they did not find a strong relationship between SOI and stream flow, possibly because PDSI is a mathematical function of temperature and precipitation and provides a general indication of drought, whereas stream flow tends to integrate climatic processes over interseasonal time scales, and this seasonal averaging may limit forecast accuracy. Second, Rajagopalan et al. (2000) found correlation between summer PDSI and winter Pacific Ocean sea surface temperature anomalies (Niño-3 index). However, they also found epochal variations in this correlation, with the period of 1963-1995 showing weaker teleconnections than the period 1895-1962. Of course, without a means of predicting these epochal shifts in teleconnections, such variation tends to confound statistical forecasting methods based on the entire historical record.

Correlations between various periods of SOI, NAO, and basin-aggregated stream flow, along with stream flow autocorrelation coefficients, are shown in Table 1. (Concurrent correlations between average annual stream flow and SOI and NAO were found to be –0.28 and –0.24, both significant at the .10 level.) Highlighted values indicate observations that could provide annual stream flow forecasts for water contract decisions made by the LCRA each November. For instance, January-March SOI values appear the best choice to predict stream flow in the following year.

Table 1. Indicator-streamflow correlation coefficients.

	Streamflow	
SOI	**Jan-Dec(0)**	**Jan-Dec(1)**
Jan(0)-Dec(0)	**-0.2763**	0.0603
Jan(0)-Mar(0)	**-0.2324**	**0.2477**
Jan(0)-Apr(0)	**-0.2563**	**0.2401**
Jan(0)-May(0)	**-0.2801**	0.2130
Jan(0)-Jun(0)	**-0.2932**	0.2062
Jan(0)-Oct(0)	**-0.2779**	0.1069
Apr(0)-Jun(0)	**-0.2649**	0.1004
Apr(0)-Sep(0)	**-0.2285**	0.0268
Jun(0)-Nov(0)	-0.2149	-0.0213
Jul(0)-Dec(0)	-0.2005	-0.0827

	Streamflow	
NAO	**Jan-Dec(0)**	**Jan-Dec(1)**
Jan(0)-Dec(0)	**-0.2364**	0.0117
Jan(0)-Mar(0)	-0.0469	0.0007
Apr(0)-Jun(0)	**-0.2891**	0.1046
Jul(0)-Sep(0)	-0.0852	**-0.2180**
Oct(0)-Dec(0)	-0.1074	0.1636
Jan(0)-Sep(0)	**-0.2290**	-0.0912
Jun(0)-Dec(0)	**-0.2293**	-0.0150
Jul(0)-Nov(0)	-0.1503	-0.0472
Jul(0)-Dec(0)	-0.1431	-0.0419
Aug(0)-Oct(0)	-0.2027	-0.1264

	Streamflow
Flow	**Jan-Dec(1)**
Oct(0)	0.1625
Nov(0)	**0.2372**
Dec(0)	**0.4429**

It is widely hypothesized that interdecadal North Pacific variability modulates ENSO-precipitation teleconnections (e.g., Gershunov and Barnett, 1998), but Rajagopalan et al. (2000) were not able to conclude that PDO has any effect on ENSO-precipitation teleconnections in central Texas. Our analysis indicated that concurrent values of the PDO index and annual streamflow had a correlation of 0.34, significant at the .05 level, but there were no significant lagged correlations to indicate forecast value. Previous studies (*e.g.*, Hamlet and Lettenmeier, 1999) have found modest dry or wet trends to be in association with long periods of each PDO phase. An apparent recent shift in the PDO to the cool phase could cause a general trend towards drier than normal conditions in the Southwest U.S., whereas this region has for the past 25 years been experiencing wetter than normal conditions. Although the PDO index alone has only a very slight correlation with streamflow in Texas, the combined effect of PDO and ENSO may be much more significant. Our analysis indicated that years of warm ENSO (El Niño) and PDO were followed by lower than average streamflow (72% of historical mean), whereas years in which the cool ENSO phase (La Niña) coincided with cool PDO were followed by significantly higher than average flow (133% of historical mean). Such information is potentially helpful during the repetition of such coincidental events, but verification of predictive skill is limited due to their infrequent occurrence. During the 60 years of historical record, only five sets of warm phases coincided (1952, 1964, 1966, 1969, 1995), and four sets of cool phases occurred (1939, 1984, 1985, 1986).

Along with the various teleconnections, we have begun exploring the potential of soil moisture as a climate indicator and predictor of streamflow over the LCRA. The VIC Retrospective Land Surface Data Set (Maurer et al., 2002) has been used to obtain estimates of fractional soil moisture over the LCRA region for the time period of 1950-1999. A great deal of effort went into selectively extracting the NetCDF format data from the VIC Retrospective Land Surface Data Set over the LCRA and importing it into ArcView GIS. The first step of the analysis was to perform a streamflow-runoff comparison between the VIC Retrospective Land Surface Data Set runoff and observed streamflow in the LCRA region. The modeled annual aggregate runoff and observed annual aggregate streamflow were found to track each other quite well with a correlation of 0.75. This assurance that the VIC Retrospective Land Surface Data Set performs reasonably well in predicting trends in runoff compared to observations gave us confidence in using the VIC modeled soil moisture for correlation analysis with observed streamflow. Visual comparison between trends in VIC modeled soil moisture and observed streamflow (see Figure 1) showed the potential of soil moisture as a climate indicator and predictor of streamflow.

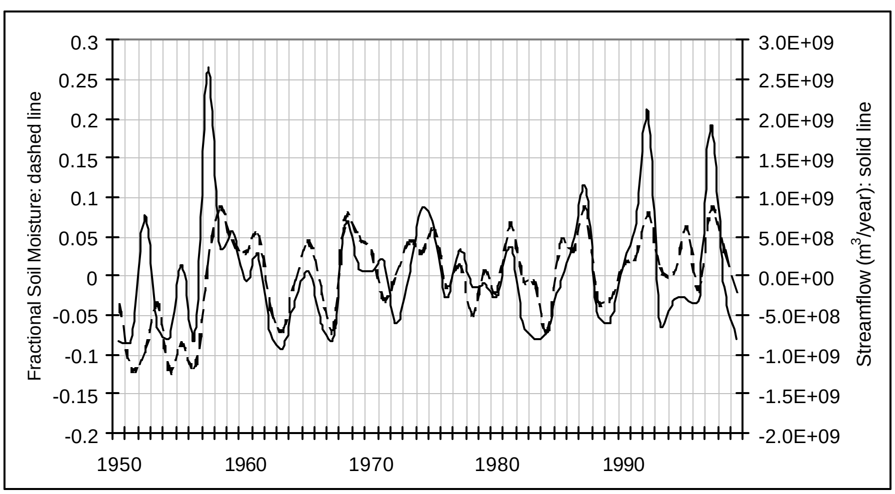

Figure 1: Climatology plot of annual average fractional soil moisture anomalies and annual aggregate observed streamflow anomalies. Anomalies defined as deviations from temporal average.

Correlation between average fractional soil moisture and aggregate observed streamflow was computed for different combinations of annual and seasonal time periods and lag times. The aggregation and averaging was performed over the LCRA region and for the specified time period. Correlations between VIC modeled soil moisture and observed streamflow at a 0-season (for seasonal aggregates) or 0-year (for annual aggregates) lag were found to be in the range of 0.5~0.7. Figure 2 shows the correlation between annual aggregate observed streamflow and various average fractional soil moisture at 0-year lag. Although not shown below, the correlations decreased rapidly as the lead-time between the soil moisture and streamflow increased. At the 1-season and 1-year lag, the correlations decreased below significance. The demonstrated need for coincident year (0-lag) soil moisture and streamflow to get significant correlations highlights the importance of soil moisture forecasts. For example, a reliable prediction of seasonal or annual average soil moisture for next year has potential as a climate indicator and predictor of next year's annual aggregate streamflow.

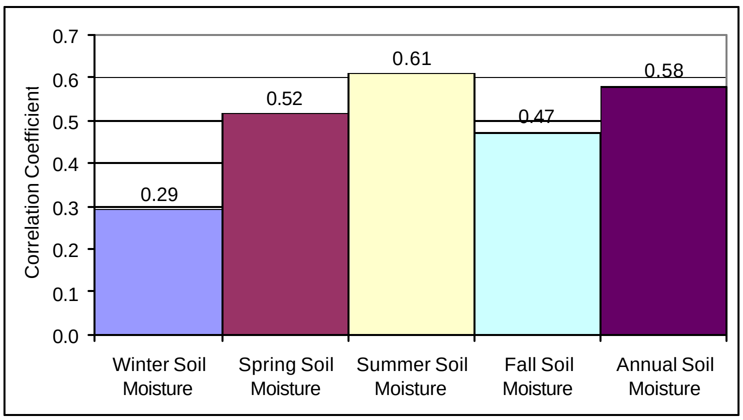

Figure 2: Correlation between annual aggregate observed streamflow and various average fractional soil moisture for coincident (0-lag) years.

Derivation of Maximal Skill Forecasts

Upon identification of potential climate predictors, the next step is to select a predictor or combination of predictors that provide the best forecasts (as measured by an appropriate skill score). The procedure followed in this study was similar to that of Piechota and Dracup (1999): Use discriminant analysis coupled with Bayesian updating to derive probabilistic categorical forecasts, with forecast skill measured by the Ranked Probability Score (RPS). An intermediate step in the discriminant analysis was to apply kernel density estimation to derive probability density functions for each flow category (High, Medium, Low) and each climate predictor (e.g., SOI, NAO). Results for SOI and NAO are shown in Figure 3.

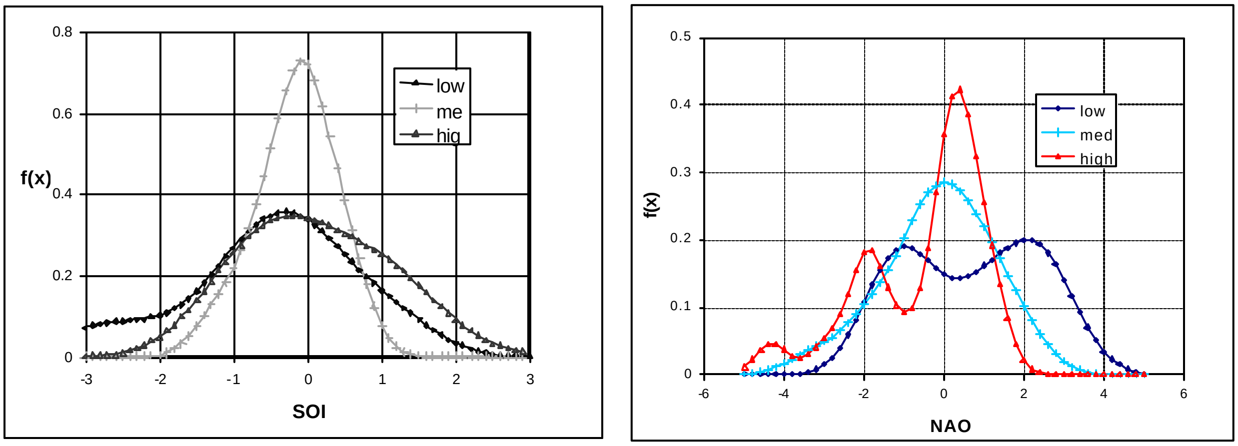

Figure 3. Conditional probability density functions for each flow regime.

Kernel density estimation indicated some apparent shifts in probability of occurrence, but also yielded some inconclusive results. For instance, as seen in Figure 3, high SOI values (greater than 1.0) appear more likely to precede high flows than medium or low flows. However, high SOI values also tend to precede low flows more frequently than medium flows, perhaps indicating that high SOI values correspond to highly variable conditions in the subsequent year. Similarly inconclusive results were obtained for NAO and stream flow persistence. However, as shown in Figure 3, high NAO values (greater than 2.0) appear much more likely to precede low flows than medium or high flow, indicating that NAO may serve as a predictor of drought.

The next step is to apply Bayes Theorem to develop posterior probabilities of each flow category conditioned on the observed indicator values. Cross validation can then be performed, and forecast skill can be evaluated using the Ranked Probability Score (RPS) and Ranked Probability Skill Score (RPSS). Results for three predictors are shown in Table 2, where RPSS indicates the relative improvement of using a forecast over climatology alone. These results indicate that NAO is the only one of the three predictors with skill, providing a 7.9% improvement in RPS values over climatology. However, while NAO appears to be a good predictor of low flows, it is a very poor predictor of average flow conditions. SOI is the best predictor of medium flows (as might be expected from the pdfs shown in Figure 3), but surprisingly has no skill as a predictor of low or high flow conditions. Based on these results, stream flow persistence does not appear useful as a predictor of annual flow. (However, stream flow persistence does appear useful as a predictor with lead times of 3-6 months, with 4-6% improvement in RPS values over climatology, which will be considered in future work.)

Table 2. Ranked Probability Scores and Skill Scores for annual flow prediction.

	RPS			RPSS		
	SOI	**NAO**	**OCT**	**SOI**	**NAO**	**OCT**
L	0.295	0.216	0.313	-0.063	0.223	-0.126
M	0.102	0.144	0.103	0.084	-0.293	0.069
H	0.279	0.254	0.263	-0.004	0.084	0.052
Ave	0.225	0.205	0.226	-0.014	0.079	-0.019

Some preliminary analysis has been done to determine if improved forecasts can be derived from combinations of indicators. As described by O'Connell (2002), the weighting method of Piechota and Dracup (1999) led to inconclusive results. Data mining approaches (e.g., Steinberg and Cardell, 1998) are now being explored and appear promising.

Generation of Stream Flow Ensembles

In the absence of a calibrated and verified hydrologic model for continuous simulation of the Lower Colorado River basin, either historic stream flow sequences or synthetic stream flow sequences from a multivariate ARMA model have been proposed for use in the development of probabilistic hydrology outlooks. Details are given in O'Connell (2002). The final step needed to incorporate climate forecasts in a scenario-based decision support model (Watkins et al., 2000) is to condition the streamflow sequence probabilities appropriately. The method currently proposed involves a simple categorical shift, referred to as the Croley-Wilks method (Croley, 2000; Wilks, 2001). This method simply adjusts the probability of flow in each regime by dividing the probability conditional to a predictor by the number of events in that regime. Although straightforward, this approach has been criticized by Stedinger and Kim (2002) because it leads to discontinuities between categories in the probability distribution, which can result in a loss of resolution or information. Stedinger and Kim recommend a distribution-oriented approach that assigns continuously shifted probabilities to each point rather than discrete categorical shifts. Adapting such an approach for nonparametric density functions will be investigated in future phases of this work.

Plan of Work in Next 3 Months

Work planned for the remainder of Year 1 includes the development and application of data mining methods to derive composite climate forecasts (i.e., stream flow forecasts based on multiple predictors). Soil moisture extracted from the VIC Retrospective Land Surface Data Set will be included in the analysis as a potential predictor. We will continue to use the Ranked Probability Skill Score to evaluate the potential benefits of these forecasts, and additional analysis will be done to estimate the statistical significance of these skill scores using bootstrap methods (Efron, 1982).

Collaboration with the Lower Colorado River Authority and technology transfer will continue with a visit by the PI to the LCRA in late Summer or early Fall 2003. Due to the untimely death of Dr. Quentin Martin, whose foresight and enthusiastic support made this work possible, Dr. Jobaid Kabir will be the contact person at the LCRA for Year 2 of the project. Dr. Kabir has committed the LCRA to continued collaboration with the PIs, and he is directly supervising the implementation of the decision support system for which climate forecasts are being derived.

In August 2003, the project will produce its first MS graduate. Mohammed Mahmoud has been working under the direction of Dr. Nykanen exploring the potential of soil moisture as a climate indicator and predictor of streamflow over the LCRA using the VIC Retrospective Land Surface Data Set. It is anticipated that Mohammed will defend his MS thesis work in early August. The project has also supported the research of a PhD student, Wenge Wei, under the direction of Dr. Watkins. Wenge has been making good progress on probabilistic forecast validation and, over the summer months, he will continue working on the generation and validation of forecasts based on multiple predictors.

Plan of Work in Year 2

Currently, there are no significant changes expected for the Year 2 work plan. The remaining tasks will be completed as follows in order to estimate the potential economic benefits of forecasts:

- Modification and application of an existing stochastic optimization model for reservoir operations (Watkins and Kabir)
- Inference of forecast-based operating rules from the optimization results (Watkins and Kabir)
- Simulation of rules derived with and without seasonal forecast information to evaluate the benefits of forecasts (Watkins, Nykanen, and Kabir).
- Evaluate usefulness of NOAA CPC forecasts of soil moisture anomalies in predicting streamflow for the LCRA region (Nykanen).

Results from Year 1 emphasize the fact that hydrologic forecasting for the water supply operations of the LCRA poses several serious challenges, including relatively short hydro-climatological records, the need for long lead times (9-12 months), the lack of a hydrologic predictor such as snowpack, and the need to consider both Atlantic and Pacific climate anomalies. Although correlations between NAO, ENSO, and PDO and stream flow are weak, it is possible that stronger correlations exist with global sea surface temperatures, and additional analysis will be done to confirm this. It is also expected that using a combination of predictors will provide modest improvements in forecast skill. However, significant improvement may require closer analysis of the hydrologic processes that contribute to stream flow, which tends to integrate climatic processes over interseasonal time scales. For instance, streamflow is a function of both surface runoff and groundwater discharge, and groundwater recharge and discharge processes often exhibit lag times markedly longer than those of rainfall-runoff processes. Furthermore, groundwater basins seldom align directly with surface watersheds, which may confound statistical analyses of climate and streamflow variables measured at specific gage locations. Spectral analysis techniques (e.g., Shun and Duffy, 1999) will be applied to address these concerns.

Correlation analysis between soil moisture from the VIC Retrospective Land Surface Data Set and observed streamflow demonstrated the need for coincident year (0-lag) soil moisture and streamflow to obtain significant correlation. Thus, if soil moisture is to be used as a climate indicator and predictor of annual aggregate streamflow, then reliable forecasts of seasonal to annual soil moisture are needed. Research during Year 2 will include incorporation of NOAA CPC forecasts of seasonal soil moisture anomalies in the LCRA Decision Support System. It is expected that this exploratory research will provide guidance to the operational research community on the skill and lead time required for soil moisture forecasts to be useful in water resources management applications.

References

Croley II, T.E. (2000). *Using Meteorological Probability Forecasts in Operational Hydrology*. Reston, VA: ASCE Press.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: SIAM.

Hamlet, A. F., and D. P. Lettenmeier (1999). "Columbia River streamflow forecasting based on ENSO and PDO climate signals." *J. Water Resour. Plng. & Mgmt.* 125(6), 333-341.

Hamlet, A. F., D. Huppert, and D.P. Lettenmaier (2002). "Economic Value of Long-Lead Streamflow Forecasts for Columbia River Hydropower." *J. Water Resour. Plng. & Mgmt*. 128(2), 91-101.

Hurrell, J. (2002, June 26). *North Atlantic Oscillation*. Retrieved August 30, 2002, from National Center for Atmospheric Research. http://www.cgd.ucar.edu/~jhurrell /nao.html

Jones, P. (2001, March). SOI data. Retrieved November 2001, from Climate Research Unit. http://www.cru.uea.ac.uk/cru/data/soi.htm